DOCUMENT RESUME

ED 034 776 24 TE 001 652

AUTHOR	McGrew, Jean B.
TITLE	An Experiment to Assess the Effectiveness of the Dictation Machine as an Aid to Teachers in Evaluation and Improvement of Student Composition. Final Report.
INSTITUTION	Lincoln Public Schools, Nebr.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C. Bureau of Research.
BUREAU NO	BR-9-F-002
PUB DATE	30 Jun 69
GRANT	OEG-6-9-009002-0061(010)
NOTE	23p.
AVAILABLE FROM	ERIC Clearinghouse on the Teaching of English, 508 So. Sixth St., Champaign, Ill. 61820 (on loan only)
EDRS PRICE	EDRS Price MF-$0.25 HC Not Available from EDRS.
DESCRIPTORS	*Composition (Literary), Educational Research, *Evaluation Methods, Evaluation Techniques, Grading, *Student Evaluation, Student Improvement, *Tape Recorders

ABSTRACT

Research on using a dictaphone for evaluation of student papers was carried out to test the technique's thoroughness, effectiveness, and capability for individualized instruction. Two classes, one an experimental and one a control group, from each of grades 9, 10, 11, and 12 were used, and each student wrote nine papers. The papers of the control groups were evaluated by marginal comments, whereas the compositions of the experimental groups were evaluated with the aid of the dictating machine. The teacher's remarks were transcribed in duplicate; one copy was returned with the student's paper while one was retained by the teacher. The first and last papers of all groups were evaluated by a three-man team in terms of content, mechanics, diction, and expression. The experimental group was found to have made more improvement than the control group on 19 of 25 comparisons. The results, too indeterminate to indicate superiority for the experimental process, suggest that the experimental procedure has merit for improving composition and that further research over a longer period of time should be undertaken. [Not available in hard copy due to marginal legibility of original document.] (LH)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

BR 9-F-002
PA 24
OE/BR

FINAL REPORT

Project No. 9-F-002
Grant No. OEG-6-9-009002-0061(010)

ED034776

AN EXPERIMENT TO ASSESS THE EFFECTIVENESS OF THE DICTATION MACHINE AS AN AID TO TEACHERS IN EVALUATION AND IMPROVEMENT OF STUDENT COMPOSITION.

Dr. Jean B. McGrew

Lincoln East Junior-Senior High School
Lincoln Public Schools
Lincoln, Nebraska

30 June 1969

The research reported herein was performed pursuant to a grant with the Office of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgement in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

TE001652

U.S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

TABLE OF CONTENTS

Summary	1
Introduction	
Background	2
Intent of Research	3
Description of Technique	3
Advantages of Dictaphone Technique	4
Survey of Literature	5
Methods	
Research Methods (Teachers)	7
Methods of Evaluating Research (Evaluators)	8
Results	
Key to Data Finding	8
Findings	9
Improvement Scores by Grade and ability level	9
Break-down of Data	10
Comparison by Categories	11
General Observations	12
Conclusions and Recommendations	
Conclusions	13
Recommendations	13
Addendum	13
Appendix	
Appendix A	14
Appendix B	16
Appendix C	17
Appendix D	21

SUMMARY

It is generally agreed that to improve composition, students should write often and have their writing carefully evaluated. English teachers, however, often find it quite difficult to do in-depth evaluations of student compositions. They simply do not have the time to do it. This research was not an attempt to provide additional time to the English teacher, but to examine a technique which might allow the teacher to be more thorough and more effective in composition evaluation in the time presently available. The technique included the use of the dictaphone as an aid to the teacher for the evaluation of student compositions and was based on the knowledge that people speak faster than they write. The study also suggested that the dictaphone technique should also allow the teacher more opportunity to spot individual problems, observe individual progress, and to individualize assignments to composition students.

Two classes from Lincoln East High School, one an experimental group and the other a control group, were selected from each grade level 9, 10, 11, and 12. The compositions of the experimental group were evaluated with the aid of dictating equipment, i.e. comments on each paper were read into a Dictaphone Dicta-Mite. These comments were transcribed and the typed transcription appended to the paper returned to the students. The students in control class had their papers evaluated in a traditional fashion, i.e. pencil notations in the margins. Each group wrote a total of nine papers. The first and last papers were compared by a team of three evaluators from public schools other than Lincoln East High School and the University of Nebraska. The evaluators compared the second paper with the first paper in terms of individual improvement in the categories of Content, Mechanics, Diction, and Expression. The scores assigned by the evaluators were averaged and a comparison between the control and experimental group made.

Although the experimental group scored higher on 19 of the 25 comparisons made, the difference was not sufficiently large to claim superiority for the experimental process.

This experiment was originally proposed to last a full school year. The fact that it was conducted in one semester undoubtedly influenced the results. The data gathered, while not conclusive, does suggest that the experimental procedure does have merit for improving composition and should be pursued over a more extended period of time.

Background and basic assumptions for research. English is the one subject in school which enrolls a majority of students at nearly every grade level. This majority enrollment is understandable. The English language, written or oral, is the basic tool of communication in our country and is thus necessary, or at least relevant, to the study of every other subject in the school curriculum. In spite of this intensive exposure to English, there is still considerable opinion that many students do not write as well as they should or could. While most of these opinions are subjective and impressionistic, they do include enough professional concensus to warrant concern among educators and to justify efforts toward improving this skill throughout the school system. It might be asked then, what kind of program should or could be developed in a school that would improve writing skills? The precise nature of such a program would undoubtedly be debated by professionals in the field. There are, however, certain conditions generally accepted by educators as conducive to any learning situation which would certainly be applicable to any writing improvement program. These conditions are listed below:

1. Students should be taught by teachers who are well qualified in their fields.
2. Class learning is irregular, i.e. not all students in a class learn at the same rate nor do they all have the same difficulties with the material to be learned. Thus, instruction should be individualized to the greatest degree possible.
3. Instruction should not only be individualized, but should allow for continuous evaluation over as long a period of time as possible.
4. In order to create individualization of instruction and to develop means of continuous evaluation, teachers should be allowed as much time and auxiliary help as the budget of the school will allow.

Most conscientious administrators try to arrange the school program to maximize these conditions. There are limits, however, on just what can be done. For example, individualized instruction and continuous individual evaluation are very demanding of the teacher's time. Improvement in either of these areas requires an increase in the amount of time available to teachers. This is usually done by (1) a reduction of the teacher's class load or (2) providing the teacher with para-professional assistants. This is difficult for some school systems since any major reduction of class load or the addition of more than a few para-professionals is often financially prohibitive. For such school systems other less expensive means need to be found if any additional time is to be provided the teacher. While it is true that the provision of additional time would be helpful to a teacher in any area, it is particularly helpful where writing improvement is desired. Students should be given frequent opportunities to write and have their writings carefully examined by the writing instructor. It is this careful and continuous evaluation, however, which requires an inordinate amount of the instructor's time.

Intent of Research. It was not the intent of this research to suggest a means by which additional hours could be put at the disposal of the writing instructor. It was, rather, intended to suggest a technique by which the time currently available could be better utilized to increase individualization of instruction and continuous evaluation. The key to this technique involves the use of the dictaphone by the writing instructor in the evaluation of students'papers, and is based on the common knowledge that people can speak many times faster than they can write.

Description of Technique. A description of the technique used is as follows:

1. Students would be asked to number the lines of those writing assignments the teacher expected to examine and evaluate. Example below:

    1. When, in the course of human events, it becomes
    2. necessary for one people to dissolve the political bands
    3. which have conected them with another, and to assume,
    4. among the powers of the earth, the separate and equal station
    5. to which the laws of nature and of natures
    6. god entitled them, a decent respect to the opinion's of mankind,

2. The teacher, while evaluating the paper, would simply speak observations or corrections into the dictaphone. In examining the example above, the teacher would speak the following into the dictaphone:

    "line three - you need another n in connected"
    "line five - apostrophe in nature's"
    "line six - capital G in God"
    "line six - no apostrophe in opinions

    The teacher might make other more general comments as well. For example:

    "line twelve - good introductory sentence"
    "line fifteen - verb does not agree with subject"
    "line twenty - this sentence is awkward. Get to the point"

    The teacher might also wish to make an extended comment at some convenient point such as the end of the paper. For example:

    "John, your use of alliteration is excellent, but you use it too often. You need to get right to the point in this paragraph. It tells the reader what your story is all about."

    At the end of the paper, the teacher would summarize her comments for each of several categories, which would be standardized among all the teachers using the dictaphones, and would include the following:

1. Mechanics 2. Appropriateness of Diction 3. Accuracy and Effectiveness of Expression 4. Content 5. Total Effect 6. Grade (These categories suggested by Dr. Frank Rice, co-director, Nebraska Curriculum Development Center)

Finally, the teacher would turn over the dictaphone cylinders to a typist. It should be noted that at this point the demands on the time of the professional teacher are complete. The typist would transcribe the teacher's evaluations of each student's paper into one original and one carbon copy. The original would be stapled to the student's paper and the carbon would be returned to the teacher and filed under the student's name.

Advantages of Dictaphone Technique. Advantage #1. Since people speak many times faster than they write, the teacher would be able to speak criticisms of the student's writing much faster than she could write the same criticism. Thus, within a set amount of time, the teacher should be able to greatly expand the criticism and comment on each of the student's papers as compared to the traditional method. The teacher with the dictaphone should be able to do considerably more in evaluation than the teacher writing out all criticism if each allowed ten minutes per paper. Advantage #2. Using the dictaphone technique, one copy of the teacher's evaluation is stapled to the student's paper and returned to the student. Under the conventional method of paper grading, the student's paper is returned with various corrections and comments written in appropriate places. It is hoped by the instructor that the student will learn by carefully studying these comments. The student, however, may simply file the paper away in his notebook. When this happens, the teacher's comments have not contributed much to the learning situation. In using the dictaphone technique, however, the student's paper is returned unmarked. All comments and corrections are on the appended typed page. Thus, the teacher may require the student to make the suggested corrections or changes on the unmarked paper. This is easily checked, especially if the student is required to use an uncommon color, such as red or green, in the corrections. The important thing, though, is that the student reads the teacher's comments and physically makes the corrections on his own paper rather than reading (or not reading) the teacher's comments already on his paper under conventional techniques. Advantage #3. Under the conventional system of evaluating student writing, all of the teacher's comments are written on the student's paper. If the teacher keeps the corrected paper the student does not benefit from the teacher's comments and corrections. This problem is resolved, however, using the dictaphone technique. The typist simply transcribes the dictaphone belt into an original and one copy. The student receives the original copy appended to his paper and the teacher receives a copy which may be placed in the student's file. The value of both the teacher and the student having copies of the teacher's comments is self-evident. The student needs a copy to facilitate self-improvement. The carbon copy provided the teacher serves an equally important function. By evaluating each paper according to established categories, i.e. mech-

anics, diction, accuracy of expression, content and total effect, and by filing these evaluations in chronological order, the teacher can easily see what patterns develop over a period of time.

Furthermore, by having this evidence in a categorical and sequential arrangement, the teacher should also be better able to individualize instruction. For example, the teacher might note that a student showed marked improvement in spelling over a period of time but was erratic in the use of grammar.

The above examples are not intended to suggest that the same could not be done without the use of the dictaphone, but that the dictaphone technique easily provides a continuous record of the teacher's evaluations. This is especially conducive to a longitudinal appraisal and the individualization of instruction without any additional expenditure of time to the teacher.

Survey of Literature. Very little is found in the professional literature regarding the use of dictation machines in composition evaluation. There are, however, a few articles that should be noted.

Robert Lumsden described a project in the November, 1961, issue of the English Journal[1] which is quite similar to the one included in this research. The teacher's criticism was dictated into the dictation machine, transcribed and a copy of the criticism appended to the student's paper. Lumsden indicated that this procedure did appear to save time:

> Teacher X processed forty-four papers in 350 minutes by hand. She did forty-four equivalent papers in 180 minutes with a Voicewriter. Teacher Y did 116 papers in 1728 minutes by hand. He did 116 equivalent papers in 705 minutes with a Voicewriter.[2]

Lumsden also offered the opinion that ". . . an English teacher benefits more from 100 hours of stenographic help than he does from 100 hours of lay reader help."[3] Lumsden implied that this process has been helpful to students, but provided no comparative data in his report to back up this opinion. Furthermore, his description seemed to indicate that only one copy of the teacher's criticism was typed and that copy was appended to the student's paper. Thus, it appears that the teacher did not have duplicate copies from which he could make longitudinal assessments and prescribe individualized assignments.

1. Robert Lumsden, "Dictation Machines as Teacher Aids," English Journal, Vol. L, No. 8, (Nov. 1961) pp. 555-556.

2. Ibid., p. 556

3. Ibid.

Bruce Hawkinson reported using a tape recorder (instead of a dictation machine) with a class of students.[1] Teacher comments and criticisms were read into the tape recorder and each student then took his turn listening to the play-back on his paper. Hawkinson noted that "The students seem to realize that comments by the instructor are more complete, since he can verbalize in one minute about six times as much as he can write in the same time . . . He will not only point out an error or weakness but he will also call attention to why it is an error. Since he can say about six times as much in a minute as he can write in the same time, he has not only the urge but also the opportunity to be positive."[2] Hawkinson indicated that 66% of the "tape graded" students raised their grades as compared to 13% of the "conventionally" graded students.[3] The report did not indicate, however, whether this grade applied only to composition or whether this included the student's grade for the English class as a whole. Here again, there appeared to be no duplicate copies for the teacher, i.e. copies from which the teacher could make longitudinal evaluations and could prescribe individual assignments.

Bernard Tanner of Cubberly High School in Palo Alto, California reported a project similar to the one described by Hawkinson, i.e. the direct playback to the student.[4] He found that time saved on one place was lost in another. He found that teachers were spending too much time monitoring machines and showing students how to use the equipment. Originally, this experiment included a typist to transcribe the teacher's recorded comments for each student. This was discontinued, however, because of expense and because ". . . the written comments which, although more nearly complete than those in marginal notes, frequently appeared in transcription as ambiguous in expression, careless in style, and faulty in punctuation, structure, and spelling."[5] It would appear that close supervision and in-service training of the teachers and typist should be able to "iron out" this kind of transcription problem.

Finally, Lee Frank Lowe of North Mercer Junior High School, Mercer Island, Washington reported using the tape recorder to correct English

1. Bruce Hawkinson, "Grading Themes with a Tape Recorder," The Education Digest, Vol. XXX, No. 7 (March, 1965), pp. 48-49. Originally printed in Educational Screen and Audiovisual Guide, XLIII (December, 1964) pp. 698-699.

2. Ibid., pp. 48-49.

3. Ibid.

4. Bernard Tanner, "Teacher to Disc to Student" The English Journal, Vol. LIII, No. 5 (May, 1964), pp. 362-363.

5. Ibid., p. 362

themes.[1] This also used the direct play-back approach. Lowe briefly described the procedure he had used and concluded that "Theme correcting has never been as worthwhile an experience as it is now."

Research Methods. During the second semester of the 1968-69 school year, one teacher was selected for the experiment in each of the 9th, 10th, 11th and 12th grades at Lincoln East High School. Each teacher, in turn, selected a representative class to act as the experimental group and a class to act as a control group. The fact that these classes were selected as a matter of convenience rather than randomly should not in any way prejudice the research since the final evaluation by outside evaluation was based on an improvement factor for each individual rather than any pre-set standard or grade. (This will be discussed in the Evaluation section). Each teacher had the experimental and control classes write a three page theme in class on a common topic selected by the teacher. The assigned topic differed among the four teachers but was the same for the control and experimental classes of each teacher. These themes (both control and experimental) were classified by the teacher into three groups--above average (designated as Red), average (designated as White) and below average (designated as Blue). This classification was marked on the top of the paper along with the grade level and the word "pre" to designate that they were pre-test papers. For example, a paper showing above average ability in composition skill written by a 12th grade student would be marked "Pre-Red-12". The reason for this initial classification (which was unknown to the evaluators) was to be able to compare the improvement factors of the three ability groups between the experimental and control classes as well as the classes as a whole. No mark was placed on the body of the theme. These papers were filed with Mrs. Shirley Doan, chairman of the English department. It was agreed among the teachers that nine equivalent themes would be assigned to both the experimental and control classes. Each third paper was to be done in class and the ninth one would constitute the comparison paper, i.e. to compare for improvement with the first paper. As the last paper was to serve as the comparison paper or post-paper, it was assigned on a topic similar to the topic assigned by the teacher on the first paper. This was done to facilitate comparison by the evaluators.

The papers of the experimental group were evaluated by means of the teacher reading her comments into a Dictaphone Dicta-Mite in a manner similar to that described in the previous section entitled Intent of Research. The control group was evaluated by conventional means, i.e. pencil notations in the margins of the papers returned to the students. Special attention was paid to four categories in both groups. These categories included CONTENT (defined as logic, coherency, material covered, and degree of understanding) MECHANICS (defined as punctuation, capitalization, spelling and paragraphing), DICTION (defined as word choices, originality) and EXPRESSION (defined as

1. Lee Frank Lowe, "Theme Correcting Via Tape Recorder," The English Journal, Vol. LII, No. 3 (March, 1963), pp. 212-214.

awkward sentences, sentence structure, wordiness, tone, attitude, perspective and redundancy). The teacher's comments spoken into the Dicta-Mite for each student paper in the experimental class were transcribed by a typist in duplicate. One copy was appended to the student's paper and returned to him. The other copy was filed in a file folder in the English department office. These files were available for review by teachers and students.

Method of Evaluating Research. As mentioned in the previous section, the last paper written (ninth) by both the students in the control and experimental groups was on a topic assigned by the teacher and similar to the topic assigned for the first paper. This allowed the evaluators to compare similar kinds of papers. The evaluators included two junior high school teachers, two senior high school teachers and two English supervisors from the University of Nebraska. None of the evaluators were on the staff of Lincoln East Junior-Senior High School. The two junior high teachers and one English supervisor evaluated the 9th and 10th grade papers (experimental and control). The two senior high teachers and the other English supervisor evaluated the 11th and 12th grade papers (experimental and control). The papers of the experimental and control groups were coded so the evaluators did not know which papers belonged to which group.

The last paper written constituted the "post" paper. Nothing was written on this paper nor was it evaluated by the teacher in any way. The last paper was stapled to the first paper to form the "pre-post" set. These were arranged in a folder alphabetically by grade. The alphabetizing mixed the control and experimental groups. All papers were then evaluated by three evaluators in terms of improvement of the second paper over the first paper. (For detailed instructions on the evaluation procedure, see Appendix A). Each evaluator placed their evaluation marks on a prepared scoring sheet. (For example of scoring sheet, see Appendix B). None of the evaluators saw the scoring sheet of another evaluator. When all scoring sheets were returned, the marks assigned by the three evaluators for each student were recorded on a 3 x 5 card (one card for each student) and averaged to produce a composite average of the improvement score. The 3 x 5 cards were then assembled by grades into the control and experimental groups for comparative analysis of the data.

Key to Data Findings. The following key will be helpful in understanding the enclosed data:

Red----above average student	1 point---no visible improvement
White--average student	3 points--some noticeable improvement
Blue---below average student	5 points--considerable improvement

Improvement Scores by Grade and Ability Level

Classification & No. of Students	Content	Mechanics	Diction	Expression	Total
		9th Control			
Red (7)	2.0	1.7	1.7	1.8	1.8
White (11)	2.2	1.7	1.5	1.7	1.8
Blue (8)	2.0	1.8	1.4	1.5	1.7
All average (26)	2.1	1.7	1.5	1.7	1.76
		9th Experimental			
Red (9)	2.0	1.8	1.9	2.0	2.0
White (7)	2.2	1.7	1.8	1.9	1.9
Blue (7)	2.1	1.9	1.7	1.9	1.9
All average (23)	2.1	1.8	1.8	1.6	1.93
		10th Control			
Red (14)	2.3	2.4	2.2	2.4	2.4
White (13)	1.7	1.9	1.7	1.9	1.7
Blue (4)	1.5	1.4	1.5	1.5	1.5
All average (31)	1.8	1.9	1.8	1.9	1.99
		10th Experimental			
Red (11)	1.6	1.5	1.5	1.5	1.5
White (10)	2.0	2.0	1.8	1.8	2.0
Blue (6)	1.9	1.7	1.7	1.6	1.7
All average (27)	1.8	1.4	1.7	1.6	1.73
		11th Control			
Red (8)	1.4	1.4	1.4	1.4	1.4
White (15)	1.4	1.3	1.1	1.2	1.2
Blue (5)	2.0	1.4	1.6	1.7	1.8
All average (28)	1.7	1.4	1.4	1.5	1.41
		11th Experimental			
Red (6)	1.6	1.0	1.5	1.4	1.3
White (15)	1.4	1.8	1.6	1.8	1.7
Blue (9)	1.3	1.5	1.2	1.2	1.4
All average (30)	1.4	1.4	1.4	1.4	1.53

Classification & No. of Students	Content	Mechanics	Diction	Expression	Total
		12th Control			
Red (9)	1.7	1.3	1.4	1.3	1.4
White (12)	1.6	1.3	1.3	1.3	1.2
Blue (7)	1.6	1.6	2.0	2.0	1.8
All average (28)	1.6	1.4	1.5	1.5	1.41
		12th Experimental			
Red (6)	2.1	1.0	1.7	1.7	1.6
White (15)	1.8	1.4	1.5	1.6	1.7
Blue (7)	2.6	1.9	2.1	2.4	2.3
All average (28)	2.3	1.5	1.9	2.0	1.82

Break-down of Data

(Asterisk used to help visual identification of higher score)

1. Comparison of <u>all</u> control groups with all experimental groups.
   A. Control (9-12) - 1.64 Experimental (9-12) - 1.74*

2. Comparison of each control group with each experimental group at <u>each grade level</u>.
   A. Control (9th) - 1.76 Experimental (9th) - 1.93*
   B. Control (10th)- 1.99* Experimental (10th)- 1.73
   C. Control (11th)- 1.41 Experimental (11th)- 1.53*
   D. Control (12th)- 1.41 Experimental (12th)- 1.82*

3. Comparison of each <u>ability category</u> in control group, i.e. "above average (Red)," "average (White)," and "below average (Blue)" with the like ability category in experimental group.
   A. "Above average" (9-12) Control --- 1.84* "Above average" (9-12) Experimental -- 1.62
   B. "Average" (9-12) Control --- 1.49 "Average" (9-12) Experimental -- 1.79*
   C. "Below average"(9-12) Control --- 1.72 "Below average"(9-12) Experimental -- 1.80*

4. Comparison of ability categories between control and experimental groups at <u>each grade level</u>.

	"Above Average"	
	<u>Control</u>	<u>Experimental</u>
Grade 9	1.8	2.0*
Grade 10	2.4*	1.5
Grade 11	1.4*	1.3
Grade 12	1.4	1.6*
Total	1.82*	1.62

(4 Continued)

"Average"

	Control	Experimental
Grade 9	1.8	1.9*
Grade 10	1.7	2.0*
Grade 11	1.3	1.7*
Grade 12	1.2	1.7*
Total	1.49	1.79*

"Below Average"

	Control	Experimental
Grade 9	1.7	1.9*
Grade 10	1.5	1.7*
Grade 11	1.8*	1.4
Grade 12	1.8	2.3*
Total	1.49	1.79*

Comparison by Categories

Category #1. Content (Logic, coherency, material covered and degree of understanding)
All Control (9-12)--1.79 All Experimental (9-12)--1.88*

Category #2. Mechanics (Punctuation, capitalization, spelling and paragraphing)
All Control (9-12)--1.60 All Experimental (9-12)--1.61*

Category #3. Diction (Word choices, originality)
All Control (9-12)--1.58 All Experimental (9-12)--1.71*

Category #4. Expression (Awkward sentences, sentence structure, wordiness, tone, attitude, perspective, redundancy)
All Control (9-12)--1.56 All Experimental (9-12)--1.64*

General Observations. The overall improvement factors of 1.64 for the control group and 1.74 for the experimental group are difficult data to interpret standing alone. There is little previous research to suggest any kind of norm for how well a student could or should improve in a semester. It is also recognized that the ability to quantify improvement scores in composition remains a subjective and uncertain procedure. However, the procedure used in this research, i.e. three separate evaluators following prescribed guidelines and a prescribed scoring system has to be considered about as reasonable method for quantifying judgment as there is presently available.

Considering a score of 1.00 as "no improvement" and a score of 3.00 as "some noticeable improvement," it would have to be observed that overall scores of 1.64 and 1.74 are closer to the "no improvement" category than they are to the "some noticeable improvement" category. Viewed from another perspective with a range of 4 points between 1 point (no improvement) and 5 points (considerable improvement) the improvement factors of 1.64 and 1.74 can be translated into 16% and 19% improvement. Here again, no norm exists to judge whether these percentages constitute average, above average or below average improvements. It would have to be admitted that the 3% advantage of the experimental group is not sufficiently significant to warrant any "break through" claims for the experimental group.

It might be suspected that the "above average" groups would show the least improvement as they had initially shown themselves to be competent writers and had "less room" for improvement. This did not prove to be the case. The improvement factor for all "above average" was 1.74 as compared with 1.63 for the average and 1.76 for below average. It is also interesting to note that the greatest improvement came from two distinctly different groups, i.e. the 10th grade above average control group (2.4) and the 12th grade below average experimental group (2.3). A slight tendency might be seen where the higher improvement did rest with the "above average" control group (1.84 to 1.62) but the improvement favored the experimental group in the "average" (1.79 to 1.49) and the "below average" (1.80 to 1.72) category. The widest advantage for any group was found in the "average" ability category where the experimental group was favored by .3. It is perhaps important to note that this group constituted 50% of all the students in the experiment.

In comparing the control and experimental groups by composition categories, i.e. Content, Mechanics, Diction, and Expression, the scores in each category all favored the experimental group but not by a significant amount. The greatest difference was found in Diction (Experimental - 1.7, Control - 1.58). The least difference was found in Mechanics (1.61 to 1.60). The highest improvement by any group favored the experimental group in the category of Content (1.88). The lowest was the control group in the category of Expression (1.56).

Conclusions and Recommendations. The data favors the experimental group in three out of the four classes participating in the experiment. The data favors the experimental group in all four composition categories, i.e. Content, Mechanics, Diction and Expression. In both cases, though, the differences are slight and not really sufficient to make any claims for the dictaphone method of evaluation based only on the data in this experiment. The data does show, however, that even though the differences are slight, they are fairly constant in favor of the experimental group, i.e. 19 of the 24 comparisons favored the experimental group.

Several things can be noted that were hinderances to the experiment and which may have influenced the outcome.

1. The original proposal called for a full year study of the control and experimental group with a week's orientation for the teacher prior to start of school. As this study was funded for only a semester, there was no time for orientation and additional time was lost waiting for the dictating machines to arrive. For someone not familiar with dictating equipment, it is quite important to have time to simply manipulate and practice with the equipment to gain maximum efficiency. This is reflected in the teachers' remarks in Appendix C.

2. Since this study was confined to one semester, the number of papers the student was asked to write and that teachers were required to grade (9) was more than the students normally do. This increased load may have caused students to view the experiment as something which increased their work load which, in turn, may have influenced their reaction to the participation.

3. An important part of the proposal called for a continuous examination by the teacher of the dictated comments kept in the student's folder (the duplicate typed sheet). Here again, time limited any extensive amount of examination and "feed back" from the folder. This, in turn, allowed for very little remedial prescription based on identified problem areas for individual students.

Recommendations. The improvement of the experimental group over the control group was not sufficiently great to warrant any supportable claims for the experimental procedure. The fact that the improvement of the experimental group was fairly consistent (19 out of 25 comparisons) leads to a suspicion--subjective though it may be--that the experimental procedure may still be a valid method of improving composition. It would seem advisable that this experimentation be continued with an orientation time provided the teachers and with at least a full year time elapse to compare the two groups. This would allow the writing assignments to be less "bunched up" and also allow some of the support actions by the teacher, e.g. individual examination of folders to prescribe individual remediation, to take effect.

ADDENDUM: (See Appendix D for per paper cost of experiment).

APPENDIX A

Evaluation of Dictaphone Experiment

The intent of this evaluation is to compare the writing improvement between two groups of students, an experimental group whose compositions were evaluated with the aid of dictating machines during the semester and a control group whose compositions were evaluated by conventional means. Each student (control and experiment) wrote a paper at the beginning of the semester, eight intervening papers during the semester which were evaluated by either conventional or experimental means, and a final paper in a similar form to the first paper. The first and last papers are stapled together, the first on top and the second on the bottom. It is the job of the evaluator to evaluate the second paper in terms of the first, i.e. what kind of improvement took place (if any) in the second paper in four defined categories when compared with the first paper. The four categories and definitions are included below:

A. Content - Logic, coherency, material covered and degree of understanding.
B. Mechanics - Punctuation, capitalization, spelling and paragraphing.
C. Diction - Word choices, originality.
D. Expression - Awkward sentences, sentence structure, wordiness, tone, attitude, perspective, redundancy.
E. Average of A - D - (The evaluator need not calculate this average. This will be done at the office.)

The evaluator will receive a folder of papers in alphabetic order and will receive a grading sheet also in the same alphabetic order. As these have to be graded by three different people, please try to maintain the order. The evaluator will be comparing the second paper of the two stapled together to the first one (top) using the following scoring criteria:

1 point - no improvement took place in the category in the second paper
3 points - some noticeable improvement took place in the category in the second paper
5 points - considerable improvement in the second paper
2 & 4 points - can be used if the evaluator feels a shading off of 1, 3, or 5 points is needed
0 points - used if negative improvement has taken place

Example: Let's say that after reading the two papers for Sam Adams the evaluator came to the following conclusions:

1. Some noticeable improvement took place in the second paper in the category of CONTENT - this would be scored as a 3 and placed in the appropriate box on score sheet.

(APPENDIX A continued)

2. Considerable improvement took place in the second paper in the category of MECHANICS - scored as a 5 in the appropriate box on the score sheet.

3. No improvement in DICTION - 1 in the box.

4. Negative improvement took place in EXPRESSION, i.e. the first paper was better in this category than the second - score as a 0.

The grade sheet would look like the one below.

Name	Content	Mechanics	Diction	Expression	Average
Sam Adams	3	5	1	0	

When the evaluator has completed the evaluation of all the papers in the folder, place the score sheet in the folder on top and write on the outside of the folder Evaluated by (your name). This will prevent us from sending the same folder back to you for another evaluation. Each evaluator will be evaluating two folders - either 11th and 12th or 9th and 10th.

When you've completed, call Dr. McGrew, 489-7121, and the second folder will be brought to you.

We appreciate your efforts with this project. We realize that $50.00 is probably a minimal payment for the work you are doing. We hope, however, that the results may give us a little better understanding on various means of improving composition.

J. McGrew
Principal-Project Director

APPENDIX B

10th Grade

NAME	A CONTENT	B MECHANICS	C DICTION	D EXPRESSION	AVERAGE A-D
Smith, Dan					
Stapleton, Greg					
Strom, Sue					
Tune, Mike					
Veskrna, Debbie					
Voboril, Judy					
Wallace, Nancy					
Welsch, Mark					
Wilkinson, Mary Jo					

APPENDIX C

RANDOM THOUGHTS ON DICTAPHONE GRADING OF COMPOSITIONS

Orientation and Psychological Block

No doubt exist in my mind that a period of orientation is necessary for the teacher who is involved in this work. One cannot give full concentration to the paper to be examined if technical skills are in the way. In addition, extended use helps to remove the apprehension of having one's voice recorded for someone else to hear. After a fairly short period of time, I felt this ceased to be a problem.

Time Involved

Naturally, the time spent per paper on the dictaphone is at first greater than the time spent on a paper graded by hand. However, as the teacher becomes more skilled and familiar with the machine, he is able to work faster.

If this is intended as a time saving device, the possibility is great. If the teacher only intends to do the same amount of (evaluation) grading that is possible by hand, after orientation, the overall time spent should be much less.

Student Benefit and Reaction

If this is intended as a student benefit device, the possibility is also great. Given the same amount of time per paper, as is generally used to grade a composition by hand, the teacher is able to offer more criticism and "in-depth" criticism than is possible when all must be written out.

Disadvantages

The fact that the teacher really needs to find a quiet room away from other people makes the method a little inconvenient. English teachers naturally have many papers to grade and they train themselves to grade them when bits and snatches of time are available and we do lose this time when using dictaphones for paper sets and this can make our evening and after school work heavier. (However, maybe this type of grading is not good, because it lacks the coherence required for a dictaphone evaluation.)

The other disadvantage, at East High, for me, with the experimental situation was that I was not able to offer as much training for student teachers in paper evaluation as I would have liked. However, I was able to work out more than I originally thought.

Overall Statement

I feel that this method of grading has many possibilities, but that other things must take place along with the institution of such a method. Time is the basic factor. If we were to use them at the present time for the sake of saving time, that would be fine; however, if our intent is to use them for student benefit purposes, then the English

teacher's load must be lightened so she can devote the necessary time without a drop in her performance in preparation, instruction, test evaluation and supervision.

Respectfully submitted,

Mrs. Ann Barry

DICTAPHONE EXPERIMENT

I am extremely interested in learning the results of the experiment; I will be whole-heartedly for this type of grading if a few of my questions are answered in the evaluative results.

Is improvement in writing related to lengthy evaluations of each paper or is it related to the natural maturing process?

Did the students in the upper one fourth of a particular class gain as much as the students in the lower one fourth of the same class? (Placing in class could be based on IQ results, grades in English class, or achievement tests.) In other words, does a teacher improve a student's writing who consistently gets a one or a student's writing who consistently falls in the six-seven category no matter how many comments and by what means? To put it bluntly, is it worth the time it takes with the dictaphone? Could the teacher's time better be spent planning and teaching? Is composition more important in an English class than literature?

Is spoken English an effective means of evaluating written English? Many times I found as I read through the transcribed material that I had corrected such things as wordy sentences with wordy sentences, illogical sentences with illogical sentences, organizational problems with just as silly organizational problems, and word choices with uninteresting word choices. Obviously, written English vocabularies and style is different from spoken and it is thus obviously difficult to grade one with another. Perhaps written comments with personal conferences would solve some of this problem and yet retain some use for the dictaphone. One could have say four out of five 5-10 minute conferences on tape and the fifth person to person.

One last perhaps seemingly minor area that disturbed me was the four areas of grading. Many times it was difficult to determine just where one was to discuss some parts of the paper.

Sherry Ehrlich

(APPENDIX C continued)

DICTAPHONE COMMENTS

The dictaphone composition experiment was an interesting approach to dictaphone grading, but I don't think it was the first of its kind. This was one factor that I thought might have been expressed in the pre-experiment meetings. The orientation (of the staff involved) into the program was generally rather inadequate. A feeling of "do or die" was held among the four involved. This might have been avoided by a more "personally interested" orientation on the part of those advocating the program.

I felt the experiment had several excellent points, however.

The personal level of the typed sheets was helpful and softly accepted by students -- no animosity on the part of any student, even those receiving low grades, was witnessed.

Personal comments from students about parental approval was heard on several different occasions. "It gives us something to work with," was the typical comment.

The fact that more could be said in less time would have to be a strong point also. This, however, was a "sore spot" throughout as the tendency to point out all errors was ever present.

One very important drawback to the machine itself was that constant repetition was needed and this became very frustrating to the grader. It made one feel that he/she did not want to go to the next set -- that he could not and still remain sane. This was something I had not experienced with a red pen in hand.

The mechanical operation of the machine was an inconvenience also. Human error in operation, weak batteries, and faulty tapes were problems. These were particularly frustrating because at times one could not tell when they had occurred and several papers could be graded without a comment recorded.

I view the possibility of its use as valuable, but only to those who readily and willingly accept it. Some may find it priceless while others find it worthless. To advocate its blanket use would be foolish.

Jim Findley

Dictaphone grading has facets that recommend, as well as discourage, its use in evaluating compositions.

One inevitable problem in using a transcription of a dictaphone

tape is the difference between verbal and written language. If students do learn by example, the English teacher becomes conscious of each word said into the dictaphone, trying to speak as she would write, a skill that at the least requires time to master. Colloquialisms or idioms acceptable in verbal communication appear less than acceptable on the typed page. To ask students to "understand" does not negate the inherent problem of the impressionable student "learning" from the transcript which was not meant to be used as a writing example.

The time spent grading with the dictaphone is often double the time needed in grading "by hand." This may be due to the teacher's efforts to speak as she writes and the four categories of evaluation which were handled separately, often necessitating reading the paper up to four times in order to evaluate each category as a whole. Due to this time factor, many English teachers already feeling the pressure of too little time to evaluate student compositions may find it prohibitive to learn to use the dictaphone when they feel quite comfortable in their own method of evaluation.

The four categories did serve the purpose of identifying very clearly what types of errors the student makes repeatedly and may help him and his teacher to pinpoint writing problems. From the carbon copy files of the transcriptions both student and teacher have concrete material with which to evaluate progress. This is obviously impossible in traditional grading methods where both teacher and student must rely on memory in student evaluation. The files also can be used for studies on problems in the class as a whole when planning a writing unit relevant to the class.

Ideally, the stenographer who types the transcript should have some knowledge of literary and composition terms to alleviate problems of misspelled or misconstrued words. To have her in the school at all times would speed the student receiving his graded paper.

Roxanne O'Gara

APPENDIX D

Cost of Typing Per Paper (Experimental Group)

1. Total cost for typing dictated comments	$294.93
2. Total number of papers typed	910
3. Cost per paper	$ .32